1 Directional Derivatives

Definition. *Given a function $f : \mathbb{R}^2 \to \mathbb{R}$ (i.e. $z = f(x, y)$), a point (x_0, y_0) and a unit vector $\overrightarrow{u} =< u_1, u_2 >$, define the directional derivative of f in the direction $\overrightarrow{u}$ at (x_0, y_0), denoted $D_{\overrightarrow{u}} f(x_0, y_0)$, to be*

$$D_{\overrightarrow{u}} f(x_0, y_0) = \lim_{t \to 0} \frac{f(x_0 + tu_1, y_0 + tu_2) - f(x_0, y_0)}{t}$$

(if this limit exists).

A directional derivative measures the rate of change of f in the direction indicated by $\overrightarrow{u}$ at the point (x_0, y_0). In other words, look at the point (x_0, y_0) in the $xy-$plane and draw the vector $\overrightarrow{u}$ starting at (x_0, y_0). Look at the whole line in the $xy-$plane along this vector. Imagine you sliced through the graph of f straight down, through this line. You get a curve; the directional derivative is the slope of the tangent line to this curve when $x = x_0, y = y_0$.

We have seen two examples of directional derivatives, namely partial derivatives:

- Suppose $\overrightarrow{u} =< 1, 0 >$. Then

$$D_{\overrightarrow{u}} f(x_0, y_0) = \lim_{t \to 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} = f_x(x_0, y_0).$$

- Suppose $\overrightarrow{u} =< 0, 1 >$. Then

$$D_{\overrightarrow{u}} f(x_0, y_0) = \lim_{t \to 0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t} = f_y(x_0, y_0).$$

The following theorem shows you how to calculate directional derivatives:

Theorem. *Let $f : \mathbb{R}^2 \to \mathbb{R}$. If (x_0, y_0) and $\overrightarrow{u} =< u_1, u_2 >$ are such that $D_{\overrightarrow{u}} f(x_0, y_0)$ exists, then*

$$D_{\overrightarrow{u}} f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle \cdot \overrightarrow{u}.$$

Proof: Define the function $t \mapsto (x, y)$ by

$$x(t) = x_0 + tu_1, y(t) = y_0 + tu_2.$$

Now we have (from the definition of directional derivative)

$$D_{\overrightarrow{u}} f(x_0, y_0) = \lim_{t \to 0} \frac{f(x(t), y(t)) - f(x(0), y(0))}{t}.$$

Notice that this is the derivative $\frac{df}{dt}$ of the function $f(x(t), y(t))$ at $t = 0$. So by a Chain Rule from the previous section, we have

$$D_{\overrightarrow{u}} f(x_0, y_0) = \left.\frac{df}{dt}\right|_{t=0} = \left.\left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\right)\right|_{t=0}.$$

However, we can explicitly calculate $\frac{dx}{dt}$ and $\frac{dy}{dt}$ since we know the formulas for $x(t)$ and $y(t)$. Also, using the formulas for $x(t)$ and $y(t)$ we know at $t = 0$ that $x(0) = x_0$ and $y(0) = y_0$. We have:

$$D_{\overrightarrow{u}} f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0)u_1 + \frac{\partial f}{\partial y}(x_0, y_0)u_2\right) = \langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle \cdot \langle u_1, u_2\rangle$$

so the theorem is proved. ■

Example: Let $f(x, y) = 4x^2y^3 - 2xy$. Find the directional derivative of f in the direction $< 1, -1 >$ at the point $(2, 1)$.
Solution: First find the partial derivatives:

$$f_x(x, y) = 8xy^3 - 2y \qquad f_y(x, y) = 12x^2y^2 - 2x$$

Next, you need a unit vector $\overrightarrow{u}$ in the direction of $\langle 1, -1\rangle$. To do this, take $\langle 1, -1\rangle$ and divide by its length (which is $\sqrt{2}$). You get

$$\overrightarrow{u} = \frac{< 1, -1 >}{\sqrt{2}} = \left\langle \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right\rangle.$$

By the theorem,

$$\begin{aligned} D_{\overrightarrow{u}} f(2, 1) &= \langle f_x(2, 1), f_y(2, 1)\rangle \cdot \left\langle \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right\rangle \\ &= \left\langle 8(2)(1^3) - 2(1), 12(2^2)(1^2) - 2(1)\right\rangle \cdot \left\langle \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right\rangle \\ &= \langle 14, 46\rangle \cdot \left\langle \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right\rangle \\ &= \frac{14 - 46}{\sqrt{2}} = -16\sqrt{2}. \end{aligned}$$

You can also define directional derivatives for functions of three (or more) variables in a similar way:

Definition. *Given a function $f : \mathbb{R}^3 \to \mathbb{R}$ (i.e. the function looks like $f(x, y, z)$), a point (x_0, y_0, z_0) and a vector $\overrightarrow{u} =< u_1, u_2, u_3 >$, define the* <u>*directional derivative of f in the direction $\overrightarrow{u}$ at (x_0, y_0, z_0)*</u>*, denoted $D_{\overrightarrow{u}} f(x_0, y_0, z_0)$, to be*

$$D_{\overrightarrow{u}} f(x_0, y_0, z_0) = \lim_{t\to 0} \frac{f(x_0 + tu_1, y_0 + tu_2, z_0 + tu_3) - f(x_0, y_0, z_0)}{t}$$

(if this limit exists).

The same theorem works:

Theorem. *Let $f : \mathbb{R}^3 \to \mathbb{R}$. If (x_0, y_0, z_0) and $\overrightarrow{u} =< u_1, u_2, u_3 >$ are such that $D_{\overrightarrow{u}} f(x_0, y_0, z_0)$ exists, then*

$$D_{\overrightarrow{u}} f(x_0, y_0, z_0) = \langle f_x(x_0, y_0, z_0), f_y(x_0, y_0, z_0), f_z(x_0, y_0, z_0)\rangle \cdot \overrightarrow{u}.$$

2 Gradients

Notice that in the two theorems above we saw the following expressions appear:

- $\langle f_x(x, y), f_y(x, y)\rangle$ (for functions of two variables)
- $\langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle$ (for functions of three variables)

Vectors of this type are important for many reasons, so we give them a name:

Definition. *Given a function $f : \mathbb{R}^2 \to \mathbb{R}$, define the gradient of f, denoted ∇f or grad f, to be*

$$\nabla f = \langle f_x, f_y\rangle .$$

The gradient is a vector-valued function of x and y. In particular $\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$.

Definition. *Given a function $f : \mathbb{R}^3 \to \mathbb{R}$, define the gradient of f, denoted ∇f or grad f, to be*

$$\nabla f = \langle f_x, f_y, f_z\rangle .$$

The gradient is a vector-valued function of x, y and z. In particular $\nabla f : \mathbb{R}^3 \to \mathbb{R}^3$.

Example: Find the gradient of $f(x, y, z) = e^{x \sin y \cos z}$.
Solution:

$$\begin{aligned} \nabla f &= \langle f_x, f_y, f_z\rangle \\ &= \left\langle \sin y \cos z e^{x \sin y \cos z}, x \cos y \cos z e^{x \sin y \cos z}, -x \sin y \sin z e^{x \sin y \cos z} \right\rangle . \end{aligned}$$

The two theorems in the previous section can be restated as one result in terms of gradients:

Theorem. *Let f be a function of several variables (2, 3 or any number of variables). Then the directional derivative of f in the direction $\overrightarrow{u}$ is given by*

$$D_{\overrightarrow{u}} f = \nabla f \cdot \overrightarrow{u}.$$

Here are some other reasons why the gradient of a function is important. I give two theorems below for functions of two variables; the theorems also hold for functions of three (or more) variables.

Theorem. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables (i.e $z = f(x, y)$. Fix a value of z, say $z = z_0$ and consider the level curve for f at height z_0. This level curve has equation $f(x, y) = z_0$. For any point (x_0, y_0) on this level curve, ∇f is normal to the level curve at (x_0, y_0).*

Proof: Let $\overrightarrow{u}$ be a unit vector tangent to the level curve $f(x, y) = z_0$ at (x_0, y_0). Then $D_{\overrightarrow{u}} f(x_0, y_0) = 0$ since the height of the function $f(x, y)$ does not change in the direction of the level curve. By the theorem in the previous section,

$$0 = D_{\overrightarrow{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \overrightarrow{u}.$$

Therefore ∇f is orthogonal to the tangent vector $\overrightarrow{u}$ at (x_0, y_0), so it is normal to the level curve. This proves the theorem. ■

Theorem. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables (i.e $z = f(x, y)$ and let (x_0, y_0) be some point in the domain of f. Then:*

- *The direction in which the value of f increases most rapidly from the point $(x_0, y_0, f(x_0, y_0))$ is ∇f.*
- *The direction in which the value of f decreases most rapidly from the point $(x_0, y_0, f(x_0, y_0))$ is $-\nabla f$.*

Proof: Notice that the direction $\overrightarrow{u}$ which gives the largest rate of increase in the value f is the direction $\overrightarrow{u}$ which maximizes the directional derivative $D_{\overrightarrow{u}} f(x_0, y_0)$. But

$$\begin{aligned} D_{\overrightarrow{u}} f(x_0, y_0) &= \nabla f(x_0, y_0) \cdot \overrightarrow{u} \\ &= ||\nabla f(x_0, y_0)|| \, ||\overrightarrow{u}|| \cos\theta \text{ where } \theta \text{ is the angle between } \overrightarrow{u} \text{ and } \nabla f \\ &= ||\nabla f(x_0, y_0)|| \cos\theta \text{ since } \overrightarrow{u} \text{ is a unit vector} \end{aligned}$$

To maximize this, we choose $\theta = 0$ so that $\cos\theta = 1$. This is tantamount to choosing $\overrightarrow{u}$ in the same direction of $\nabla f(x_0, y_0)$. (To minimize this choose $\theta = \pi$ so $\cos\theta = -1$, i.e. choose $\overrightarrow{u}$ in the opposite direction of $\nabla f(x_0, y_0)$.) ■

Example: Let $f(x, y) = \frac{x}{x^2+y^2}$. Find a normal vector to the level curve $f(x, y) = .5$ at the point $(1, 1)$.
Solution: This is asking for $\nabla f(1, 1)$; the answer is $\langle f_x(1, 1), f_y(1, 1)\rangle =< 0, -2 >$.

Example: The temperature of a point (x, y) on a metal plate is given by $T(x, y) = 100 - 2x^2 + 3y^2$. If a heat-seeking particle is dropped on the plate at the point $(2, -4)$, find the direction the particle moves in (when it first starts to move).
Solution: The particle will move in the direction that the temperature increases most rapidly, i.e. in the direction of the gradient. This question is asking for $\nabla f(2, -4)$ which is $\langle f_x(2, -4), f_y(2, -4)\rangle = \langle -8, -24\rangle$.